

# Gourmet in Bristol

Applied Data Science Capstone  
Alex Aklson, Ph.D., Data Scientist  
Amodio Savio

# Agenda / Topics

- Project Description
- Project Methodology
- Key Findings/Results
- Research drilldown
- Conclusion

# Project Description

## Objective

- A famous chef of Italian Cuisine wants to open a 5 stars restaurant in Bristol and he wants to collaborate with nearby restaurant of different cuisine.

# Procedure/Methodology

- Procedure
  - collecting data from several resources
  - define mean / percentile of available data
  - apply clustering to the list of neighbourhoods
- Key assumptions
  - The restaurants taken in account have price 3,4

# Supporting content

[data.opendatasoft.com](http://data.opendatasoft.com)

- Wards polygons
- Lsoa polygons

Wikipedia/  
Wikimedia

- Bristol  
Neighbourhoods

[Plumplot.co.uk](http://Plumplot.co.uk)

- Average House  
Price in Bristol

[Doogal.co.uk](http://Doogal.co.uk)

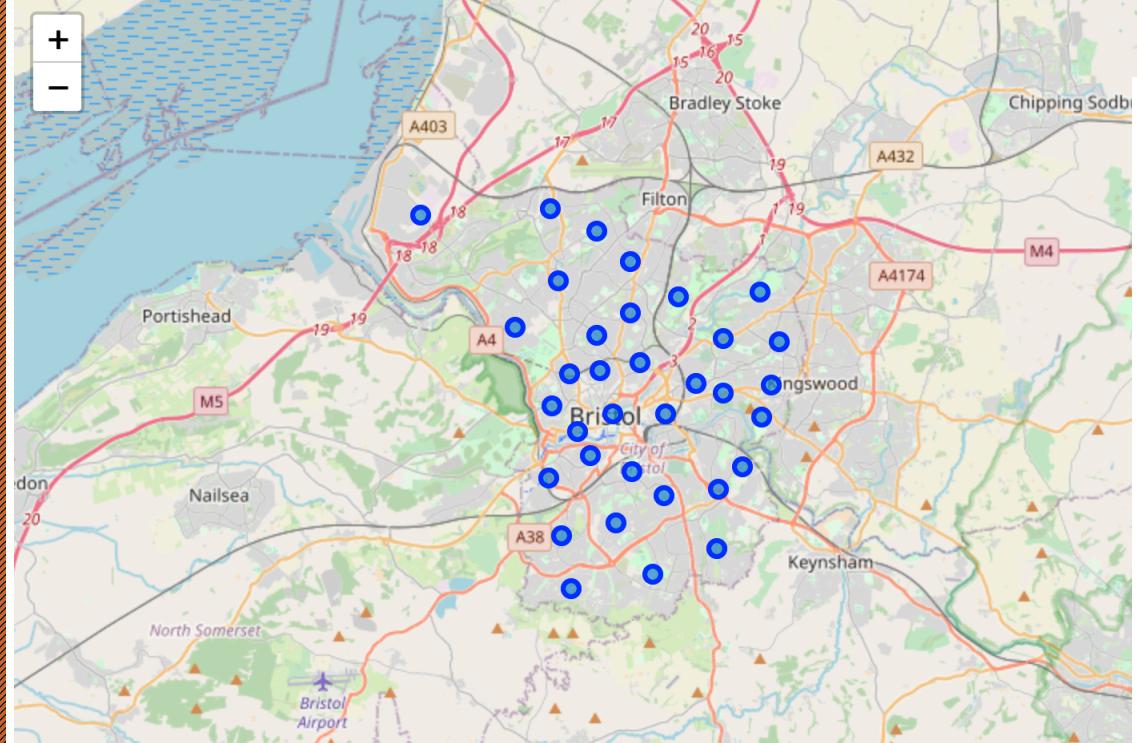
- Wards
- Lsoa
- Postcode
- Index  
deprivation
- Average Income

Foursquare

- Restaurants

# Data observations

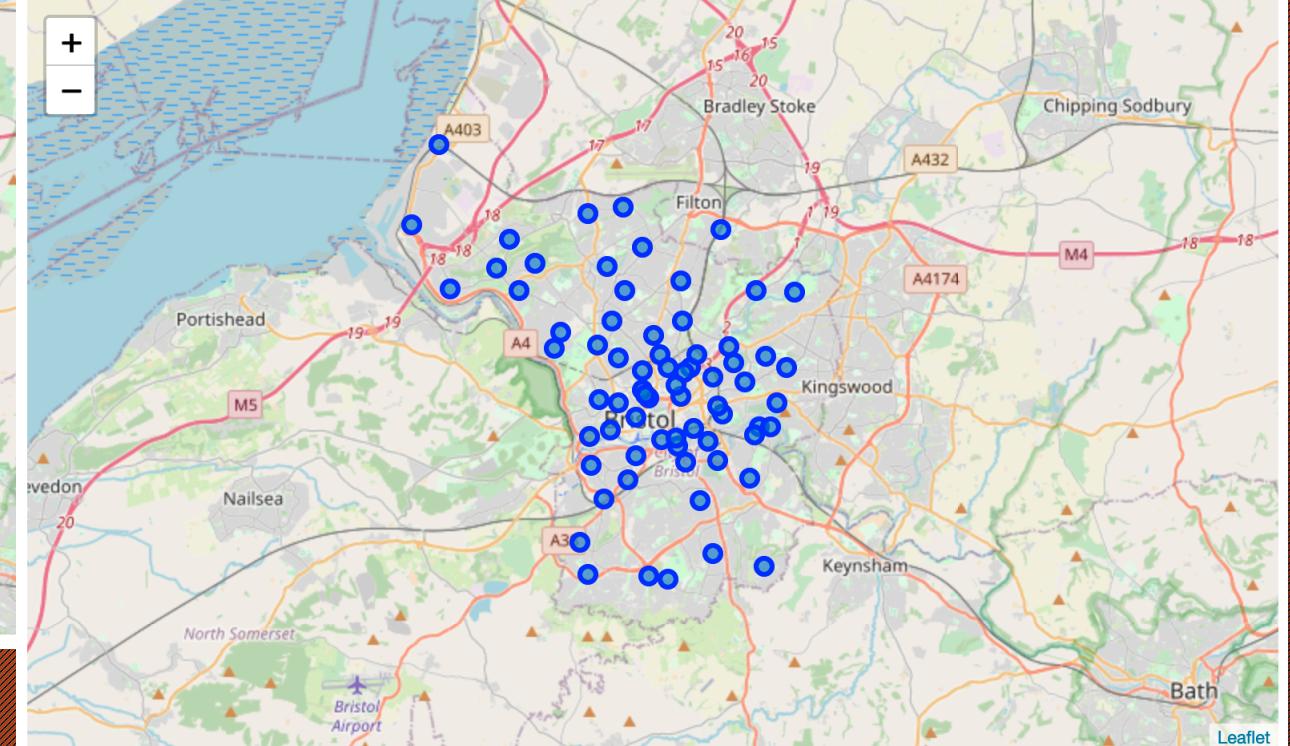
The geographical coordinate of Bristol are 51.4538022, -2.5972985.



Ward centroids (~30)

Neighbourhood centroids (~70)

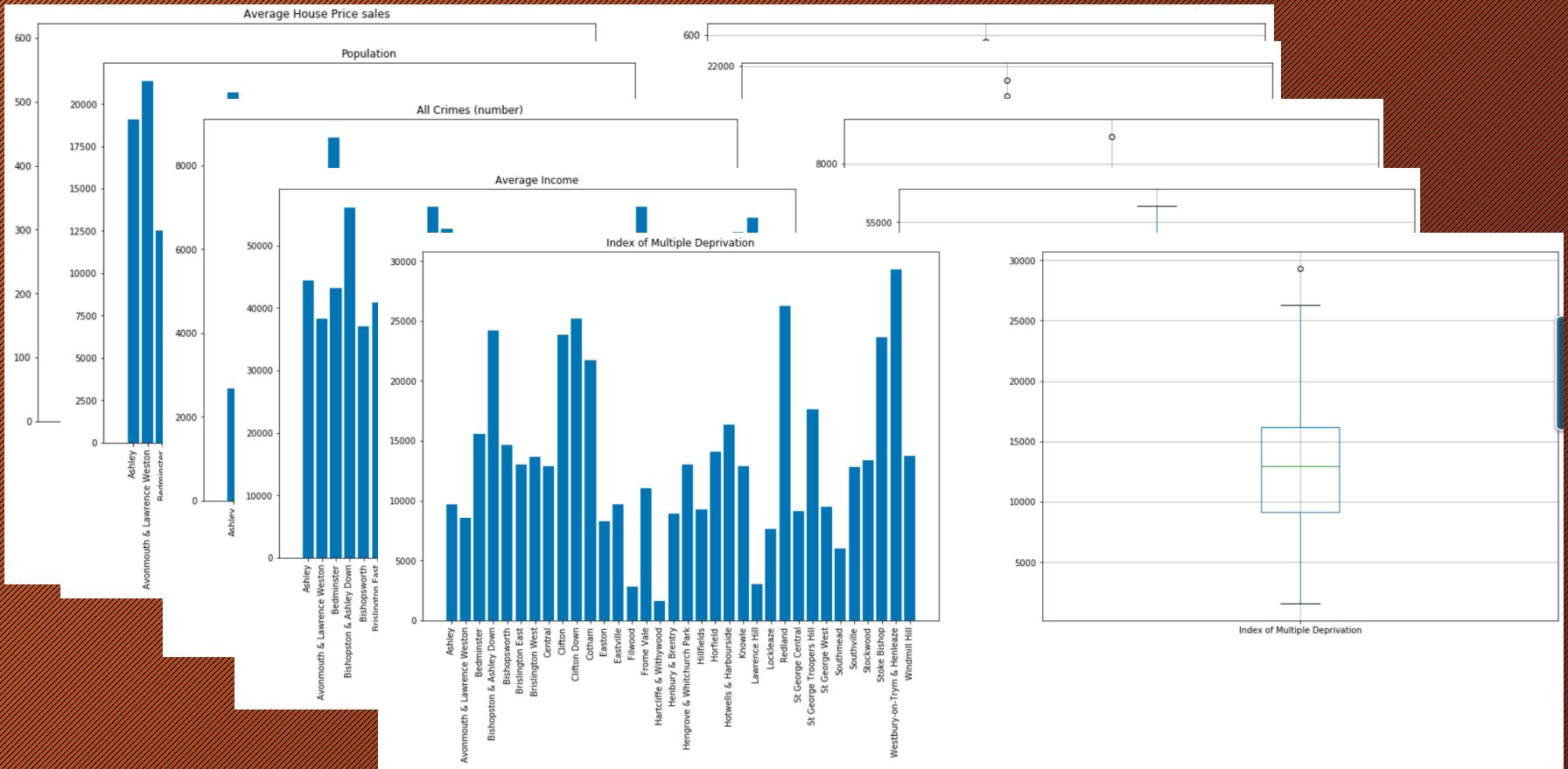
The geographical coordinate of Bristol are 51.4538022, -2.5972985.



# Data observations

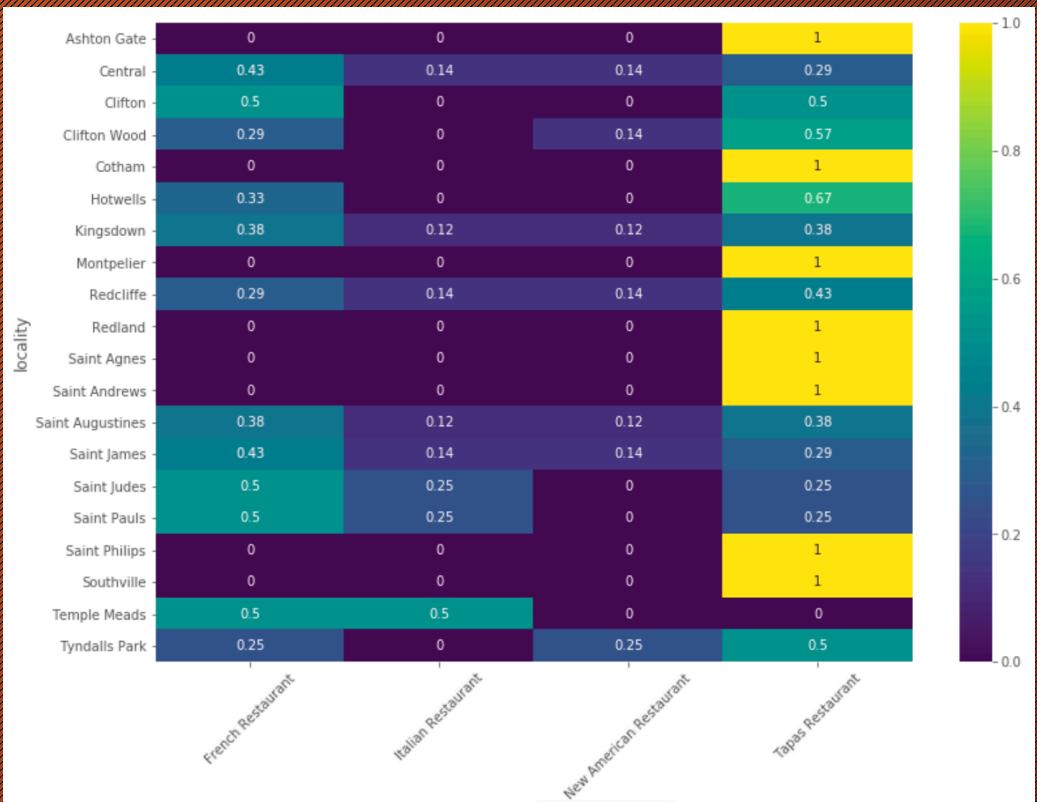
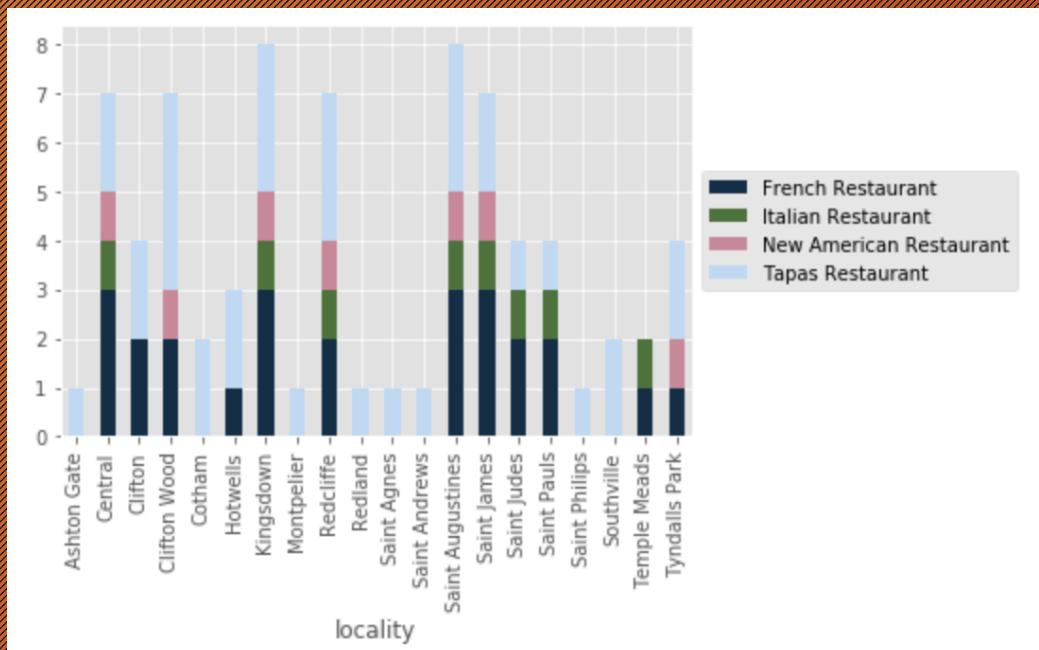
Main characteristics take in account for each ward:

1. *Average of House price*
2. *Crimes*
3. *Average Income*
4. *Population*
5. *Index of deprivation*



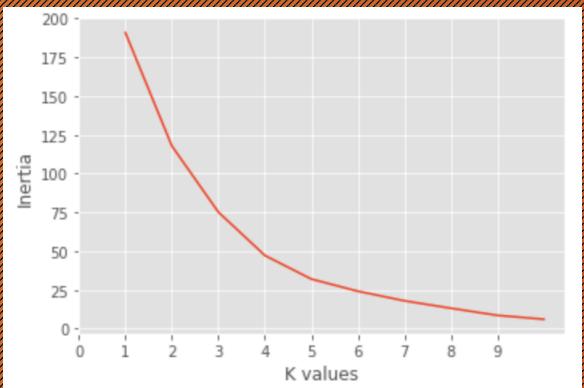
# Data observations

After fetching the restaurant from Foursquare (specific filter price = 3,4) for each neighbourhood

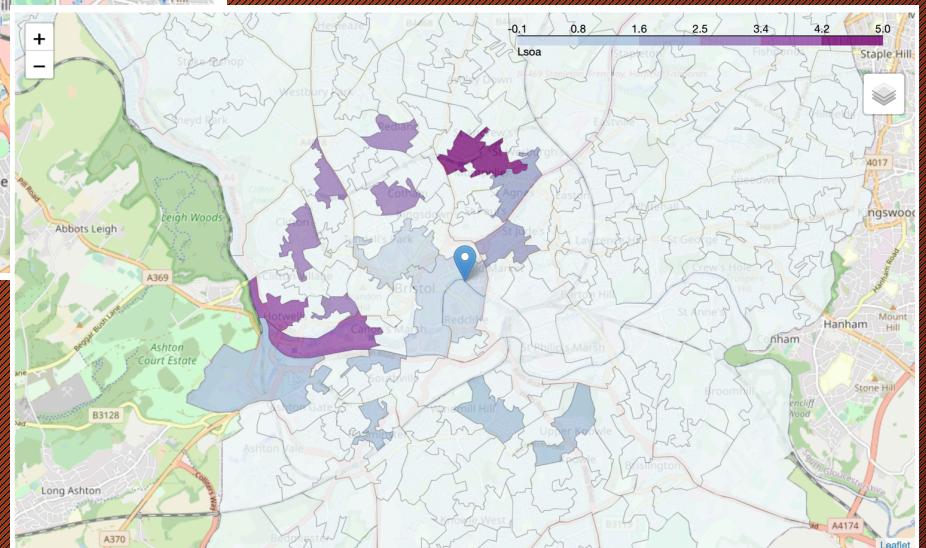
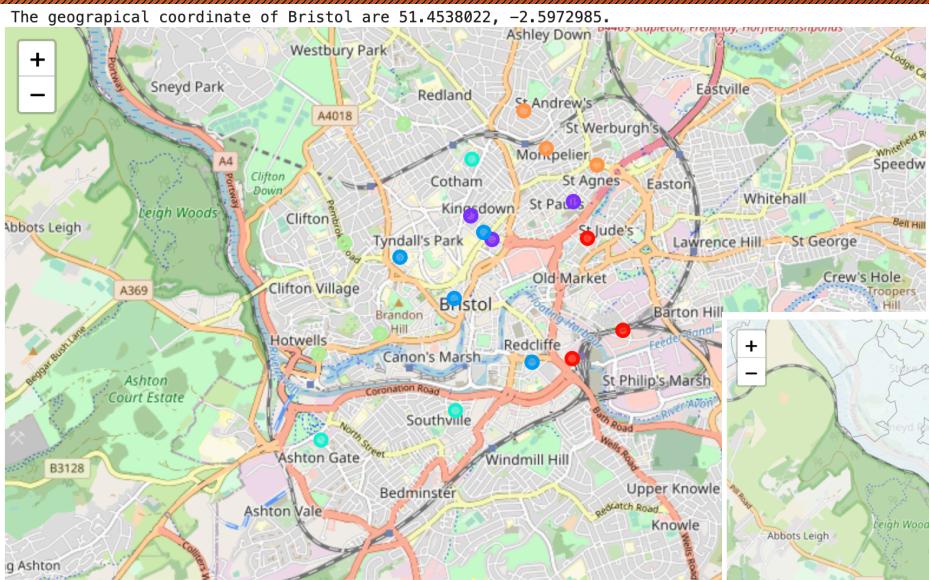


# Data observations

## Clustering with KMeans



$k = 6$  as trade off value for identifying the clusters



# Data observations

## Clusters:

1. One of characteristic of the cluster 1 is to gather neighbourhood with high population but with high crimes and most of them have highly diversity cuisine.
2. The cluster 2 has a bad diversity of restaurant but the crimes are very low.
3. The cluster 3 has neighbourhoods highly populated and the deprivation is high.
4. The cluster 4 highlights neighbourhoods quite rich and the quality of live is high: low crime, very good income, low deprivation, the population is in the average, but only one of them it's interesting *Cliftonwood* because the only one with high restaurant diversity and additional there is not competitor for our Italian Chef.
5. The cluster 5 is highly populated and house / income / deprivation are in the average, low crime but only few diversity restaurant.
6. The cluster 6 has basically only one type of cuisine *Tapas Restaurant*, the population is low, good in crimes and the average for the rest of variables.

# Data observations

## Cluster n.4

	<b>name</b>	<b>Chinese Restaurant</b>	<b>French Restaurant</b>	<b>Indian Restaurant</b>	<b>Italian Restaurant</b>	<b>New American Restaurant</b>	<b>Tapas Restaurant</b>	<b>Diversity Restaurant Ranges</b>	<b>Population Ranges</b>	<b>Crimes Ranges</b>	<b>Avg. House Price £k Ranges</b>	<b>Average Income Ranges</b>	<b>Index of Multiple Deprivation Ranges</b>
<b>5</b>	Clifton	0.0	0.33	0.0	0.0	0.00	0.67	2	3	5	5	4	5
<b>6</b>	Cliftonwood	0.0	0.29	0.0	0.0	0.14	0.57	4	3	5	5	4	5
<b>7</b>	Cotham	0.0	0.00	0.0	0.0	0.00	1.00	1	3	5	4	5	5
<b>9</b>	Kingsdown	0.0	0.00	0.0	0.0	0.00	1.00	1	2	5	4	5	5
<b>13</b>	Redland	0.0	0.00	0.0	0.0	0.00	1.00	1	3	5	5	5	5

# Conclusion

- After applying the clustering, a couple two cluster can be identified as suitable for the satisfy the requirement of chef, the first one is gathering